Response letter to Review #2

Responses from the Authors are given in blue italicized text throughout: Thank you for your very helpful feedback. Specific items of language, font size, etc. which will be addressed in the revised manuscript are colored here with gray highlighting.

General Comments:

This manuscript includes emission estimates of CO_2 and CH_4 from the Sacramento, California, area from aircraft measurements on three different days. It presents some important and interesting investigations on the sensitivity of flux estimation toward the mass balance method used, the treatment of wind measurements, the choice of background, the inclusion of entrainment at the top of the boundary layer, and different interpolation and extrapolation methods. Still, I think the manuscript needs improvement in the structure, explanation of the methods used as well as the presentation of obtained data and results. The manuscript needs major revisions before it can be accepted for publication in AMT.

The structure of the manuscript could be improved with respect to the different sensitivity studies. I recommend choosing one "best-conduct" approach, explaining and using it for the flights first, and then doing the sensitivity studies and relating their results to this "best-conduct" approach to see each choices influence on the flux calculation individually. Thus there would be one section each on the mass balance method used (Gauss vs. downwind curtain), the treatment of wind measurements (raw winds vs. mass-balanced winds), the choice of background (minimum vs. average), the inclusion of entrainment at the top of the boundary layer, and different interpolation (kriging, vertical interpolation) and extrapolation methods (kriging, constant, exponential weighting function, Gaussian fit).

The structure of the manuscript will be completely reorganized based on your suggestion of a "best conduct" approach followed by variants. More calculations were performed to make our points clear in the revised version of manuscript.

The outline of the revised manuscript will be this:

1. Introduction
2. Data and Methods
  2.1 Data Collection
  2.2 Data Gridding
    2.2.1 Extrapolation to the Surface
    2.2.2 Elliptical Fit and Measurement Interpolation (Kriging Method)
    2.2.3 Kriged GHG Mixing Ratios
3. Flux calculations
  3.1 "Best Conduct" (or "base case")
    3.1.1 Wind Treatment
    3.1.2 Background Mixing Ratios
    3.1.3 Planetary Boundary Layer Height (PBLH) and Entrainment
    3.1.4 "Best Conduct" calculated fluxes
  3.2 Sensitivity Tests

3.2.1 Sensitivity of Calculated Flux to Wind Treatment
3.2.2 Sensitivity to the Choice of Background Concentrations
3.2.3 Importance of Including Vertical Mass Transfer
3.2.4 Sensitivity of Calculated Flux to Layer Thickness
3.3 Flux Uncertainties
4. Discussion and Conclusions

The presentation of measured data differs for the three flights and the two compounds CO_2 and CH_4. Please choose the same set of figures for each flight, making sure all important data used in the flux calculation (like wind speed) is shown for all flights.

We appreciate this comment. We will include the wind data in the revised manuscript (see above) and will revise our figures.

The results of using different treatment of input data are the different flux estimates. These numbers are often only named in the text. I think a tabular representation of results for each sensitivity study (or two combined studies, like in Table 1) would increase the readability of the manuscript. I really liked Table 1 and its discussion.

We are glad that Table 1 was clear and helpful. We will expand it in the revised manuscript, separate it into two tables [Table 1 (urban scale), Table 2 (local scale)], and also include a new Table 3 showing comparison with the Turnbull et al. study and bottom-up inventories.

Table 1. Urban scale fluxes over Sacramento on November 18, 2013.

Background	**Wind**		**Urban Scale (large loop)**	
			November 18, 2013	
			CO_2 ($Mt\ yr^{-1}$)	**CH_4 ($Gg\ yr^{-1}$)**
min	**Mass-balance**	**100 m layer avg**	25.6±2.6	87.1±8.7
		500 m layer avg	26.6±2.7	88.7±8.9
		Whole column avg	26.6±2.7	88.7±8.9
avg	**Mass-balance**	**100m layer avg**	25.6±2.6	87.4±8.7
		500m layer avg	26.6±2.7	89.0±8.9
		Whole column avg	26..6±2.7	89.0±8.9
min	**Raw**		3.7±0.4	13.0±1.3
avg	**Raw**		25.5±2.6	91.1±9.1

Table 2. Local scale fluxes over landfill and rice field over Sacramento on July 29, 2015.

Background	Wind		Local Scale (small loop): July 29, 2015			
			Landfill		Rice Field	
			CO_2 ($Mt\ yr^{-1}$)	CH_4 ($Gg\ yr^{-1}$)	CO_2 ($Mt\ yr^{-1}$)	CH_4 ($Gg\ yr^{-1}$)
min	Mass - balance	100 m layer avg	0.2±0.1	7.1±2.5	0.3±0.04	2.6±0.4
		500 m layer avg	0.2±0.1	7.1±2.5	0.3±0.04	2.6±0.4
		Whole column avg	0.2±0.1	6.9±2.4	0.3±0.04	2.6±0.5
avg	Mass-balance	100 m layer avg	0.2±0.1	7.1±2.5	0.2±0.04	2.6±0.4
		500 m layer avg	0.2±0.1	7.1±2.5	0.2±0.04	2.6±0.4
		Whole column avg	0.2±0.1	6.9±2.4	0.2±0.04	2.6±0.4
min	Raw		0.4±0.1	9.0±3.1	0.2±0.03	1.7±0.3
avg	Raw		0.2±0.1	7.1±2.5	0.3±0.04	2.6±0.5

Table 3. Flux estimates for the Sacramento urban area from measurements made on November 18, 2013.

		CO2 ($Mt\ yr^{-1}$)	CH4 ($Gg\ yr^{-1}$)
Whole cylinder–AJAX	(bg = min, 100m layer avg)	25.6 ± 2.6	87.1 ± 8.7
	(bg = avg, 100m layer avg)	25.6 ± 2.6	87.4 ± 8.7
Curtain –AJAX	(bg = min)	17.3 ± 1.7	64.4 ± 6.4
	(bg = avg)	8.9 ± 0.9	24.1 ± 2.4
Turnbull et al. (2011)		13.6 (with uncertainty of ~ 100%)	
Vulcan estimates for Sacramento		11.7	
CEPAM estimate for Sacramento		10.3	

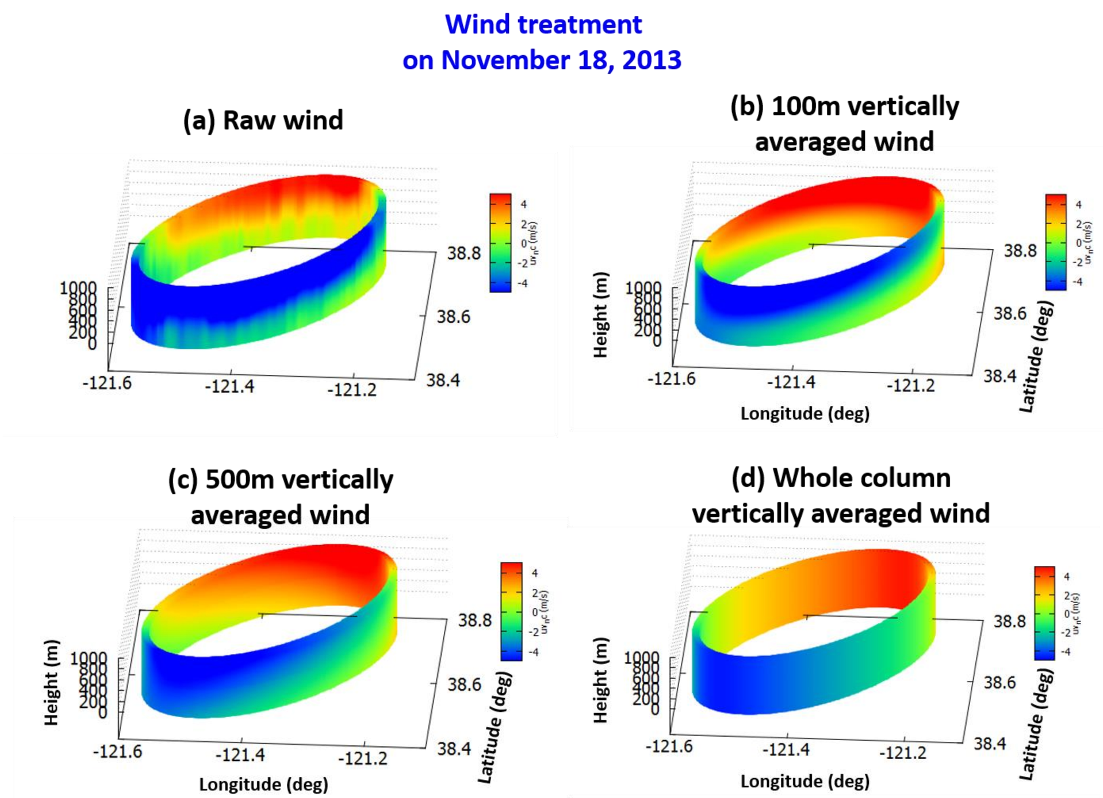
[a] Turnbull et al. (2011) data was collected in 2009; the value given here was converted from the mean reported value of 3.5 Mt C yr-1 with a 1.1% yr^{-1} increase in CO_2 flux to adjust to 2015.
[b] Bottom-up inventory estimates of the annual total emissions from Sacramento County from Vulcan (Gurney et al., 2009) and the California Air Resources Board CEPAM database (Turnbull et al, 2011) are included for comparison. The Vulcan inventory is available only for 2002, and the CEPAM database is available for 2004. We applied a 1.1% yr^{-1} increase in CO_2 flux to adjust to 2015.

I am not sure if all of your flights are well suited for flux estimation. Generally, you would look for a well-mixed boundary layer in order to decrease the uncertainty of flux below the lowest flight height. On November 17, 2015, the winds are quite different from one flight level to another and very weak at the lowest level. Furthermore, detecting the highest CO_2 concentration in an upwind part of the flight path shows that this day is not suitable for flux estimation. Also the local flux estimates on July 29, 2015, show low wind speed with changing direction.

Based on the comments of both reviewers, we have decided to remove the November 2015 case due to the vertical variability in wind speeds. However, we still believe July 29, 2015 is appropriate for testing the "closed shape" approach for flux estimates and have retained these cases in the current analysis.

Your calculation of mass-balanced wind is interesting. It seems as if the mass-balanced wind in Fig. 5 is constant with height. Shouldn't it vary with height because you use the average wind for each level? What is your surface condition? Why is kriging used? Please explain in more detail how the "mass-balanced" wind field is generated? The difference in the flux estimate does not surprise me (l.359). You have very good wind measurements, and you definitely see a change in wind speed with height in Fig. 5b. Low wind speed with high concentrations and high wind speed with low concentration might result in the same flux. Thus, Fig. 5c looks quite logical to me. If you remove all your information on the vertical wind speed change, as you did in Fig. 5e, then, of course, the flux only represents the concentration measurement. Why do you neglect your information on the wind situation? Did you ever calculate the mass-imbalance of the raw winds? Is it significant?

Thank you for your insightful discussion. Yes, the mass-balanced wind should vary with height when we use the average wind for each level. We have tested a variety of thicknesses for the vertical levels, and the results are now reported in the modified Table 1 (above). The original Fig. 5e was showing the "whole column average" case, and we apologize for not specifying this clearly in the original text. For reference, we show here the wind field with all four treatments. Panel (d) here is the same as Figure 5e in the original manuscript. We intend to include panel (b) here in an extended version of Fig. 5 in the revised manuscript.



Your use and understanding of the divergence method seems flawed. Please review the Gauss theorem and describe it correctly in l. 257. I do not believe that background is necessary for this method. You simply calculate all fluxes through the surface of the cylinder (outflow – inflow) and thus receive the change of mass within the cylinder.

We do not need to know the background value for estimating flux when adopting the circular pattern of flight. This is clearly demonstrated in our experiments, showing that the flux estimates were not sensitive to the choice of background (minimum or averaged value). We will work harder on the language in the revised manuscript.

Please check the publication of Conley et al. (2017), which nicely explains the application of Gauss theorem on aircraft mass balance flights. For your second method (the curtain downwind of the sources) you definitely need a background, and here the influence of choice of the background value on the flux estimate is quite interesting.

Thank you for the suggestion. We referenced Conley et al. (2017), and looked at their approach. We carefully considered their method and compared them to our approach.

Your calculation of the flux through the top of the cylinder is useful, but I do not understand how you determine the surface flux. First, with the Gauss theorem you need to assume that all the mass change inside the cylinder (e.g. what leaves through the surface) comes from the sources on the ground, thus what you determine is the surface flux. Second, on the ground the vertical wind speed is zero. How can the surface flux calculated with your method then be different from zero?

Since the lowest flight altitude is about ~250 meters above the ground level, we can't directly measure the surface CO_2 or CH_4 concentration or wind. The vertical velocity near the surface was very small, but it was not zero (W = -0.006 m/s for November 18, 2013 case). So we could still calculate the surface flux part. But the surface flux was much smaller than the entrainment flux, and the entrainment flux itself was not much compared to the flux computed on the "cylindrical" (wall) surface. Although the contribution of surface flux uncertainty to the total flux estimate uncertainty could be improved, we decided not to focus on this in the current study since the contribution of vertical mass transfer to the total flux estimate is relatively small.

Finally, could you calculate an overall uncertainty of your flux estimates from the different calculation methods and treatment of input parameters?

We have determined the overall uncertainty of our flux estimates based on the kriged and mass balance wind methods. By assuming that the errors of each factor are Gaussian in nature, and each measurement (e.g., CO_2 and wind) is independent (no covariance), we estimate the overall uncertainties in the calculated flux by calculate the relative uncertainties from each point to the adding the fractional uncertainties of the kriged CO_2, CH_4, and winds in quadrature, as in Nathan et al. (2015). We also consider the uncertainty due to estimate of PBLH. We will add a brief discussion to the revised manuscript comparing the distribution of the calculated fluxes reported in the revised Tables to the calculated 10% uncertainty estimate.

Please improve the consistency of your terminology. For example, you defined the two ways of treating your wind measurements as "raw wind" and "mass-balanced wind". In the following manuscript you then repeatedly use mean wind, measured wind, averaged wind, area-mean wind…
Very important points. Thank you for providing "fresh eyes" to notice this with. We should stick to using "raw wind" and "mass-balanced wind". We will improve this in the revised manuscript.

I also recommend grammar checking by an English native speaker and thorough checking of references to Figures and Sections. Furthermore, please increase the size of the axis labels and color bars on most of your plots. They are not readable.
We will make sure to improve the size and the quality of the plots in the revised manuscript.

Specific Comments and Technical Corrections:

l. 43: Which meteorological factors?
We meant the wind speed and measurement. We will rephrase it in the revised manuscript.

l. 46: "emissions fluxes" should become "emission fluxes". Please check the whole manuscript.
Thanks. We will double check this in the revised manuscript.

ll. 48-49: This should be reformulated due to the low winds on Nov. 17, 2015, and the high concentrations in the upwind part of the flight pattern.
We agree with you and have decided that this flight was not adequate for flux estimates. Hence, we will remove this flight from our analysis.

ll. 49-50: The wind variability and seasonality has not been investigated in this study. Please reformulate.
We agree with you. In the revised manuscript, we will remove those parts.

l. 51: Where do you show the influence of the distance to the emission sources?
We don't show this. We will correct the sentences in the revised manuscript.

l. 58: What is your "modeling strategy"? Do you do any modeling?
Here we meant the "statistical modeling". Using one of the geostatistical methods, kriging, we filled the gaps between measured data.

l. 61: Why don't you mention your investigation of the background and wind treatment in the Abstract?
This is a good idea. In the revised manuscript, we will mention our background and wind treatment in the abstract.
l. 65: Introduce abbreviations once and then use during the remainder of the manuscript (e.g., GHG).
Yes, we will do that in the revised manuscript.

l. 66: Is air-quality important in this study?
No, but accurate emission estimates will affect "air quality" and its regulation, so this is why we mentioned this in the introduction.

l. 69: Check your use of "give rise to"…
We will change the word to "causes".
l. 72: What is the "role of human behavior in altering the emissions"? Do we need to know it for national emission estimates?
Good points, we don't need to do know, so we will remove this part.

l. 76: What are indirect emissions?
"Indirect emission" was intended to describe information in emission databases which are inferred or extrapolated or given a time dependence that is not directly measured. We will find a more direct way to state this in the revised manuscript.

l. 78: Please give an example of a bottom-up inventory using proxy data to achieve fine spatial resolution.
Vulcan inventory CO_2 data. (Gurney et al., 2009)

ll. 83-86: What about flux estimates of European cities?
This a great point. So we will add studies for European cities in the revised manuscript: (Peylin et al, 2005; Kountouris et al., 2018).
l. 94: What do you mean with "those efforts reach general agreement on emission inventories across the cities"?
We will rephrase It in the revised manuscript. "While those efforts reach general agreement on emission inventories across the cities" will be changed to "Since current emission inventories do not consider individual characteristics of each city, they have limitations…".
l. 106: Supplementary Material
We will reword it.

l. 108: Do you really mean "uniform vertical mixing", or maybe "uniform distribution of trace gases"?
We will reword it. "uniform vertical mixing" will be changed to "uniform distribution of trace gases with altitude within the PBL and with time".

l. 110: Does this sentence ("These studies…") apply to the first or second category, or maybe both??
These studies apply to the first category.

l. 112: Does the "single-screen multi-transect method" really depend on constant wind speed? *You could also use average wind at each transect or even raw wind at each measurement point.*
Correct. The original sentence was based on the implementation of the method and the assumptions used by Karion et al. (2015). We will clarify this in the revised manuscript and then draw the comparison to our data set which contains in situ wind measurements, allowing the calculation you suggest (which is now included in the new Table 3).

l. 119: The cylinder pattern should be "around" a source and not only "near".
You're right. We will correct it in the revised manuscript.

l. 126: Why are the additional point sources considered sources of uncertainty?
If there are not included inside the oval, any uncounted point sources could be considered additional "unknown" sources of GHGs that should have been included in our urban-scale study. "Uncertainty" was probably not the best word to use, and we will fix or omit this sentence in the revised manuscript.

l. 131: I think we all got the concept of three-dimensional space (delete the parentheses).
Yes, we will delete the parentheses.
l. 134: The PBLH is not hard to measure. It is relatively easily determined from a vertical profile of temperature and humidity as you have done in this study. You even stated that the different approaches you used led to similar results. So what is difficult with respect to the PBLH? It is certainly difficult to model correctly, as you stated that substantial differences exist between models and reanalysis data. Also consider the large diurnal variability of PBLH.
We will remove "that is hard to measure".
l. 135: Please don't use "observed" in connection with models. This might confuse.
Yes, we will correct it.
l. 139: Do you really think the execution of flights is a goal of the study? Or is it merely necessary for the other goals?
One of the goals of this study is to test the flux estimate when using the cylindrical flight pattern. To reduce the confusion, we will remove the word "execute" in the revised manuscript.
l. 145: Which "value"?
We will remove this in the revised manuscript.
l. 153: Describe the three flights here and mention figure 1. It is not mentioned at all in the text.
Yes, we will do it.
l. 154: How many whole-air standards do you use for calibration?
We used two NOAA whole-air standards, a "high" and a "low". Plus we used these whole air standards to put accurate numbers onto secondary, synthetic standards, of which we have 5 or more of varying ranges. So they give us a good handle on the linearity of the instrument, across varying concentration ranges.

l. 159: Take off time is not sufficient. Please give the total flight times in UTC. Using Pacific Standard Time and Pacific Daylight Time here needs more explanation on why you give take-off times in different ways.

One flight was executed in the winter and one during Daylight Savings Time. We will re-write this section more clearly: Sampling occurred 21:10 – 22:00 UTC for the November flight (local standard time is UTC minus 8 h, 13:10 – 14:00 PST) and 20:55 – 21:45 UTC for the June flight.

Sect. 2.2: Review order of sections: I would first discuss the interpolation method and then the extrapolation.

We will consider this suggestion when constructing the revised manuscript. The significant restructuring we plan to do (shown above) may benefit from discussing the interpolation first, but we can't judge which is best until we invest the time in restructuring and re-ordering the figures.

l. 163: Reformulate: "Because the lowest flight level was typically between 250 m and 380 m above the surface …"

We will correct it as suggested.

l. 165: Sentence needs restructuring: The "unmeasured values" lead to uncertainty whether or not a "well-mixed layer assumption" is made. Split sentence!

Yes, we will work on that.

l.168: Refer to Figure 2 here. There is no reference to it in the text.

We will do that.

l. 169: What do you mean by "elevated" plume? Is it lifted of the ground or are there large enhancements of the concentration?

We mean the former: lifted off the ground.

l. 170: How exactly do you derive the background level? What is the "lowest flight measurement"? Has it got anything to do with the "lowest flight level" which you use in the formula? Is the background only determined from the lowest flight level? Why are you talking about background at this point? It is a section on extrapolation to the ground. Do you also extrapolate the background values? What do X and t stand for?

We apologize for the wording in the original manuscript at line 170. Our description of the "constant" method was unnecessarily confusing. The blue curves in Figure 2 show what we were trying to describe: at all altitudes below the yellow diamond (lowest flight level), the mixing ratio was presumed to be exactly equal to the value measured at the lowest flight level. Please also see the new figure below in supplementary materials. The bottom left panel shows more clearly how the values measured along the lowest flight level are assigned to all grid cells below the measurement altitude.

X is the given trace gas concentration, and t is the single parameter representing each point on the ellipse (eccentric anomaly)

l. 172: Do you mean that the details of the method are described in Gordon et al. (2015)?

Yes, that is what we mean.

l. 173: How is the Gaussian distribution of the plume dispersion calculated?

The Gaussian fit method is similar to the exponential fit method, except that the surface-sourced plume dispersion follows a Gaussian distribution function. For a given set of (x: height, f(x): GHG) pairs, we get the rate of change and the mixing ration at the surface ($C_{suf}(s)$) at each s parameter from the gaussian function,

$$f(x) = C_{top}(s) + (C_{suf}(s) - C_{top}(s)) * \exp(-x^2 / 2s^2)$$

Sect. 2.3: Consider renaming the section to “Measurement interpolation”.
Thank you. That is a nice suggestion.

ll.176-186: Should this be a separate section called: Projection of data to cylinder surface?
Yes, we wrestled with that. It is a bit too small to stand alone, but it might make sense to do that in the revised structure.

l. 182: What is Y?
(X(t), Y(t)) is the each point on the ellipse represented by a single parameter (t, eccentric anomaly). So X refers to the longitude and Y refers to the latitude.

l. 190: Refer to Fig. S4 as you show these differences there. Consider over plotting the measurements on the kriged and interpolated fields for better assessment of your result that kriging better captures individual plume features. What altitude range do the elliptical cylinder plots cover? Ground to PBLH? Please state in the figure caption.
Thank you for the good suggestion. To better explain the different interpolation methods, we will incorporate Fig. S4 with Fig. 2 in the revised manuscript. We plot up to the highest measurement altitude for the elliptical cylinder plots. However, for computing the actual fluxes, we only integrate the fluxes from the surface (z=0) to the top of the PBLH. We will state this clearly in the figure caption in the revised main and supplementary material.

ll. 214-227: Consider a separate section on uncertainties.
This is a good idea. We will make this as a separate section. See our new outline for the revised manuscript (page 1-2).

l. 216: Not only downwind interpolated values induce uncertainty. Upwind values as well.
You’re right. We will remove “downwind” from the sentence.
l. 226: Add “observations” behind “direction”.
Yes, we will do this in the revised manuscript.
l. 229: Remove “to the choice of background value and” because this is not the topic of this section. You do not investigate the wind characteristics but the treatment of wind measurements.
We will do it in the revised manuscript.
l. 230: Remove “In one”
We will correct it.

l. 231: What is “measured points”? How did you measure them?
We meant the discrete measurement locations (lon, lat, height) at a given time obtained by aircraft. We will endeavor to find better wording in the revised version.

l. 233: Stick to one tense (averaged, equaled).
We will be careful in using tense in the revised manuscript.

l. 233: “By assuming non-divergence, mass can be balanced.” This is correct, but is this really what you need here?

This is important. Because this is not divergent, the inflow and outflow are the same, and we can apply the mass-balance idea to our flux calculation.

l. 242: Do you assume PBLH to be constant during your flights? At what time during the flight did you measure the profile?
Yes, we assume that PBLH is constant throughout our flights. Sampling profiles occurred 21:10 – 22:00 UTC for the November flight (local standard time is UTC minus 8 h, 13:10 – 14:00 PST) and 20:55 – 21:45 UTC for the June flight.

l. 243: Is the boundary layer "growing" during your flights? How do you know?
No, we assumed that there is not sufficient time for change in the PBLH during our flights (less than 1.5 hours) We think the confusion comes from the word that we used. We will change it from "boundary layer growth" to "boundary layer height" in the revised manuscript.

l. 248: How is C(t,z) determined? How can one point surround the top of the cylinder? How is the background defined here?
We apologize for the grammar error. The sentence should have read "C(t,z) is the CO_2 concentration (g m^{-3}) at each point around the top of the cylinder (where z=h), and $C_{bg}(h)$…"
We used this formulation for each method (sensitivity test) of defining the background.

l. 252: Is the entrainment calculated from the kriged data?
Yes.

l. 263: Flux is defined through a surface. Thus it cannot be "inside" the cylinder.
You are right. Flux is defined through a surface. We will fix the grammar in this section in the revised manuscript.

ll. 265-280: See my comment in the General Comments section on the use of a background value with the Gaussian divergence theorem. If you consider inflow and outflow, you do not need a background. In your formula, the result should be invariant to the value of background mixing ratio chosen if you consider positive contributions as outflow and negative contributions as inflow.
Thank you for elaborating on this. This is also why we used the mass-balanced mean wind, so that influx mass and outflux mass are the same and the total flux estimate is not dependent on having an understanding of background mixing ratios. We will mention it in the revised manuscript.

l. 291: Use present tense.
We will change it in the revised manuscript.
l. 295: Remove "concentration".
We will change it in the revised manuscript.

l. 299: How is the kriged estimate less arbitrary in an area far away from measured values? What assumptions is it based on? Is the state of the PBL (stable/unstable) taken into account?
We believe kriging is less arbitrary because we have more constraints for formulating a kriged estimate. When we calculate fluxes based only on the measured data, without filling in the gaps between flight levels, the total flux estimate will obviously be much smaller than when we account for the entire surface of the cylinder using interpolated data. With an urban-scale cylinder (with a circumference on the order

of 100 km), it is impossible to map out the entire surface (~100 km^2) with dense measurements. Although kriging cannot be better than actual observations, it can be a good alternative to "mimic" actual data. We disagree with the reviewer's opinion that we solely rely on the kriging without an understanding of the data. We carefully performed the variogram analysis, and carefully chose the kriging parameters (sill, range, and nugget) based on the experimental and theoretical variogram obtained from the actual data we measured.
In contrast, other methods are solely based on general assumptions without the actual inspection of the existing spatial dataset.

l. 300: You don not mention the Gaussian fit method depicted in Figure 2 at all.
We mentioned it in section 2.2 (Line 172) in the original manuscript. The Gaussian fit method is similar to the exponential fit method, except that the surface-sourced plume dispersion follows a Gaussian distribution. See the explanation above. We will mention it more clearly in the revised manuscript.

Sect. 3.1: What is the influence of the different choice of interpolation and extrapolation on the flux estimate? Here a table similar to Table one would be great.
Thank you for a good comment. However, what we focus in this study is the impact of treatment of wind measurement and background on the flux estimate, not comparing different interpolation methods (although we mentioned these for completeness using Fig. 2). Furthermore, although we did show different GHG mixing ratio assumptions below the lowest flight level, we do not consider how to treat the wind below the lowest flight level. We may assume a constant wind speed and compute the flux for each of the extrapolation methods, but we are not sure how to interpret those values and we believe this will gives us additional challenges, leading to additional uncertainty in total flux estimates without understanding the physical meaning of the calculated values. Furthermore, we already mentioned that the difference of CO_2 estimate below lowest flight level could lead to the change of GHG concentration up to 20%.

l. 304: Remove "gap of the".*We will change it in the revised manuscript.*
ll. 314-320: Please mark all the locations mentioned in the text on a map so the reader can confirm your statement.
We will mark it in the revised manuscript.
l. 325: Present tense.
Yes, we will be careful about using consistent verb tense.
l. 327: Please check "a farther".
We will change it to "far".
l. 330: Maybe use the last sentence of this paragraph as its first. Good introduction.
We will restructure the sentences in the revised manuscript.

l. 350: The PBLH you determine from the vertical profile might have an uncertainty of <1%, but is this value representative for the whole measurement area with this accuracy? What about changes over time and with the location? How does a less defined PBLH influence the uncertainty?
We assumed that PBLH does not change during our 1.5 hour flight. The urban-scale area studied is approximately 20 km x 40 km with pretty uniform topography, thus we expect the PBLH to be the same throughout the sampled domain. However, we do acknowledge that a different estimate of PBLH can increase the uncertainty. Please see the response below.

Fig. S6: Looking at your method of estimating PBLH there seems to be a possible error of more than 1 % as well. In Fig. S6d it becomes clear, that you use the 50 m averaged values for checking the gradients. Then you place it at the top of the layer with the highest gradient. Here it is visible, that this point is easily 40 m above the layer where a 20 m averaged profile would see the gradient. Thus your uncertainty is around 50 m, which would be almost 10 % for a PBLH of 600 m.

That is a very good point. As you pointed out, the uncertainty of the PBLH can be up to 10% if we determine the PBLH based on the largest gradient of the vertical profile of the potential temperature. We will consider the uncertainty of the PBLH estimate and include it in the total uncertainty estimate in the revised manuscript. Based on our 3 measurements, the uncertainty due to PBLH estimate for urban scale is about ~10% , but the uncertainty due to PBLH estimate for the local-scale is about 1-5 % so that the change of PBLH does not affect the total flux estimate. As seen in Fig. S6, the vertical range of the largest gradient of potential temperature is very small, compared to the urban-scale. This leads us to another important message: the uncertainty gets larger when we deal with urban-scale flux estimate. We will include the uncertainty due to the estimate of PBLH in the total flux uncertainty estimate in the revised manuscript.

l. 355 ff: See my comment on the treatment of “mass-balanced wind” in the General Comments section.

We made new tables for the comparison as you suggested. Please see the tables (page 2-3).

Sect. 3.3: Please already refer to your Table 1 when naming the results.

Yes, we will do so.

l.267: Where is an “actual” location of the rice field? Pleas show locations on a map rather than just giving coordinates. This is very hard to visualize for a reader.

The labels in Figure S8 are awfully hard to read, and for that we apologize. We will improve them in the revised version

l. 370: “the local emissions are attributed to these high flux estimates”. Did you mean: “The high flux estimates are attributed to the local emissions”?

Yes. Thank you for pointing this out. We will fix it in the revised manuscript.

l. 374: Formulation: “mean wind vector at the dominant wind direction (positive and one direction) and speed”. How is this calculated?

Many previous studies use the mean wind averaged over the PBLH. Karion et al. (2015) estimate the total CH_4 emission in the flight region (curtain flight) using a mass balance approach. According to their study, when the mean horizontal wind speed and direction are steady during the transit of an air mass across an area, the resulting calculated horizontal flux is equal to the surface emission between the background location and the downwind measurement. This calculation required the assumption of steady horizontal wind direction, a well-developed convective PBL, and measurements sufficiently downwind of the emission source such that the emissions are vertically distributed throughout the PBL.

l. 381: There is no Table 2.

We will fix this and refer to the right one in the revised manuscript.

l. 387: Raw wind is displayed in the bottom two lines.

We will fix it in the revised manuscript.

ll. 390-391: This sentence is incomplete and not logical.

We will change it in the revised manuscript.

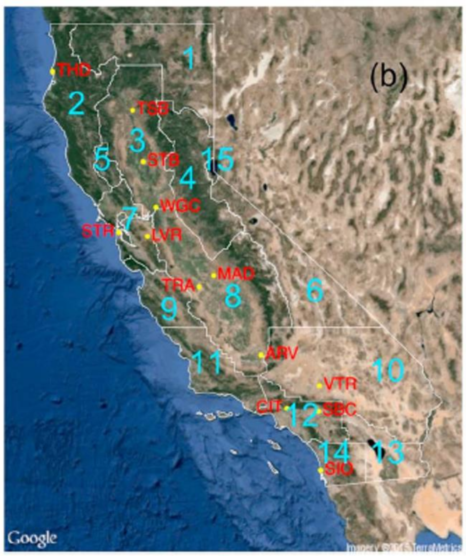
l. 394: Table 1 shows a range of 3.68 - 26.58 Mt CO_2 yr-1 for the whole city.
Thank you for catching that. It must have been a hold-over from an earlier draft.

l. 396 ff: Here you investigate the difference between using the complete ellipse and only the downwind part. This should be a separate section, and the results should be presented in another table.
Yes, we will present it in a new table (Table 3) in the revised manuscript.

l. 399: Change “From this study,…” to “According to these calculations…”
Thank you for catching that. It must have been a hold-over from an earlier draft.

l. 401: Table 1 gives a range of 13-92 Mt CO_2 yr-1 for Nov. 18, 2015.
We will correct it in the revised manuscript.

l. 402: Please indicate “Region-3” on a map.



This is from the study led by Jeung et al. (2016), and region 3 refers to the Sacramento valley. Each number represents the region classification based on California Air Basins (https://www.arb.ca.gov/ei/maps/statemap/abmap.htm).
Thus, this covers much larger area than we actually measured for the flux calculation for this study. We will explain this better in the revised version.

l. 405: Is vi) the same as i)?
We will correct that.
l. 405: Which of these does “This” refer to?
We will work on the wording in the revised manuscript.
ll. 415-422: “Note … Table 2).” All this is repetition to before and not about the topic of this section which is “vertical mass transfer”.
We will remove this part in the revised manuscript.

l. 428: Remove “First,”
We will remove this in the revised manuscript.

l. 431: Specify: “different flux calculation methods”
We will remove it.

l. 453: There is a contradiction here “the final flux estimates become similar”, because the beginning of the sentence states that the background value is a major source of uncertainty.
As you pointed out, background concentration is not important for the cylindrical flight, and we actually showed that the total flux is insensitive to the choice of background concentration when we used the mass-balanced mean wind. We stated that background value is a major source of uncertainty when we do not use mass-balanced wind for cylindrical flights. We will rewrite this more clearly in the revised manuscript.

l. 459: Insert “that” after “suggesting”.
We will insert it in the revised manuscript.

ll. 460-468: This section is a general overview of the flight results and should be placed earlier in the Conclusions.
We will move this section to the earlier in the conclusions.

l. 463: An overview of wind conditions should also be placed in the Results section.
This is a good point and we will find an appropriate place in the revised manuscript.

l. 464: This result (isolated high concentrations of CO_2) has not been shown in the Results section either.
We showed a high concentration of CO_2. This is shown in Fig. 4(c) and Fig. 5(a, d) in the original manuscript. We will more clearly discuss these plots in the revised manuscript.

l.470: Why did you expect sources to be concentrated on the downwind side?
We didn’t mean that we expect sources to be on the downwind side. What we tried to state is that horizontal flux is transported to the downwind side. We think this confusion comes from the unclear wording. We are sorry for the confusion and we will be more clear in the revised manuscript.

l. 471: “Furthermore” does not fit here.
We will work on this in the revised manuscript.

l. 471: Wind variability definitely influences the flux estimates, not only during different times of the year. So this seems logical. It would be much more interesting how large the uncertainty due to this is assumed to be.
We agree with that. We were interested in the influence of the wind treatment on the final flux estimate. However, we only used the measured wind, not any other source of wind data. It would be interesting to understand the magnitude of the uncertainty of the total flux estimate depending on the source and treatment of the wind data (e.g. measured wind vs. modeled wind with different temporal and spatial resolutions), but this is beyond the scope of this study and will be a topic of the future study.

l. 475: The size of the ellipse is another factor that appears here for the first time in the manuscript. There is no data given on how large your ellipses were and what the influence is in the Results section.

We analyzed two flight data at different size – urban scale (~ 20 x ~40 km) and local scale (< 3km). We mentioned the scale in the introduction (Line 144 in the original manuscript), and it sounds like we should reiterate it in the new Section 2.1.

l. 480 and 481: Remove two of the three "further".
We will remove this in the revised manuscript.
l. 482: Do you really want to assess: "seasonality of sensitivity of emission estimates"? Just start with the seasonality of emissions first.
We will correct it in the revised manuscript.
l. 484: Where do you show the sensitivity of emission estimate uncertainty to temperature and potential temperature?
We will revise the sentences.
ll. 490-491: This sentence needs some revision and focus.
We will work on them in the revised manuscript.

Figures
Fig. 1: There is no shading visible in Fig. 1c.
By shading we meant "color fill" in Fig. 1c. The blue color represents inflow (airflow passing through the cylinder, negative sign) and red color represents outflow (airflow passing out from the cylinder, positive sign), respectively.

Fig. 2: What is the "altitude of the lowest flight data"? Please indicate the location of these measurements on a map, giving coordinates is not very helpful.
The altitude of the lowest flight can be shown in the time series plots (shown in the response to reviewer #1). We will include this in the revised manuscript. We will also indicate the location of the measurement in the map.

Fig. 3d: Is the ellipse shown from the ground to the highest flight level or which altitude range?
Yes, the ellipse is shown from the ground to the highest flight measurement level (~ 1000 m). This is the same as shown in the time series plots (shown in response to reviewer #1).

Fig. 4: Please provide headings with the date of the flight for the left and the right column.
Fig 4 will be significantly revised.

Fig. 5: Why is the mean wind kriged? This has not been mentioned in the text.
We first kriged the measured wind and then computed the mean wind (averaged the kriged wind) at each level. We tested levels of 100m, 200m, 500m, 1000m thickness, as well as the whole cylinder.

Fig. 7: Why is there this large space between the two sets of bars? What is "area-mean"?
This just means "mass-balanced mean wind (whole vertical layer)". We will be careful and consistent with our terminology.

Table 1: Tables normally have their description above not below them.
We will modify them in the revised manuscript.

Supplementary Material:

There appear to be bits and pieces of text strewn throughout the Supplement. Please give them a heading and a number so it becomes clear where they belong, and you then may also refer to them from the main manuscript.
Fig. S1: Figure b color bar label is missing.
We will correct it.

l. 7: I am not sure if you can say that emissions are "accumulated" downwind. They are transported downwind, but accumulation would mean that there is very slow wind only.
We agree with you. We will change the wording in the revised manuscript.

ll. 11-12: This is not true. With a curtain flight it is also possible to detect emissions from more than one point source within the city, throughout the city and downwind. It gets problematic if there are sources further upwind of the city that gets mixed with the city plume and cannot be separated from it.
We will correct them in the revised manuscript.

ll. 15: You mention three types of flight patterns in the main manuscript but only show two of them here.
We will carefully examine this in the revised manuscript.

Fig S2: Reformulate "throughout the altitude". Color bar labels are missing.
We will work on them in the revised manuscript.

l. 25: "accumulated" s.a.
We will work on them in the revised manuscript.

l. 28: Why is air at lower wind speeds less dispersive?
We will remove November 17, 2015 case, so this sentence will be significantly revised.

l. 28 ff: Reformulate sentence "Both flights …"
We will remove November 17, 2015 case, so this sentence will be significantly revised.

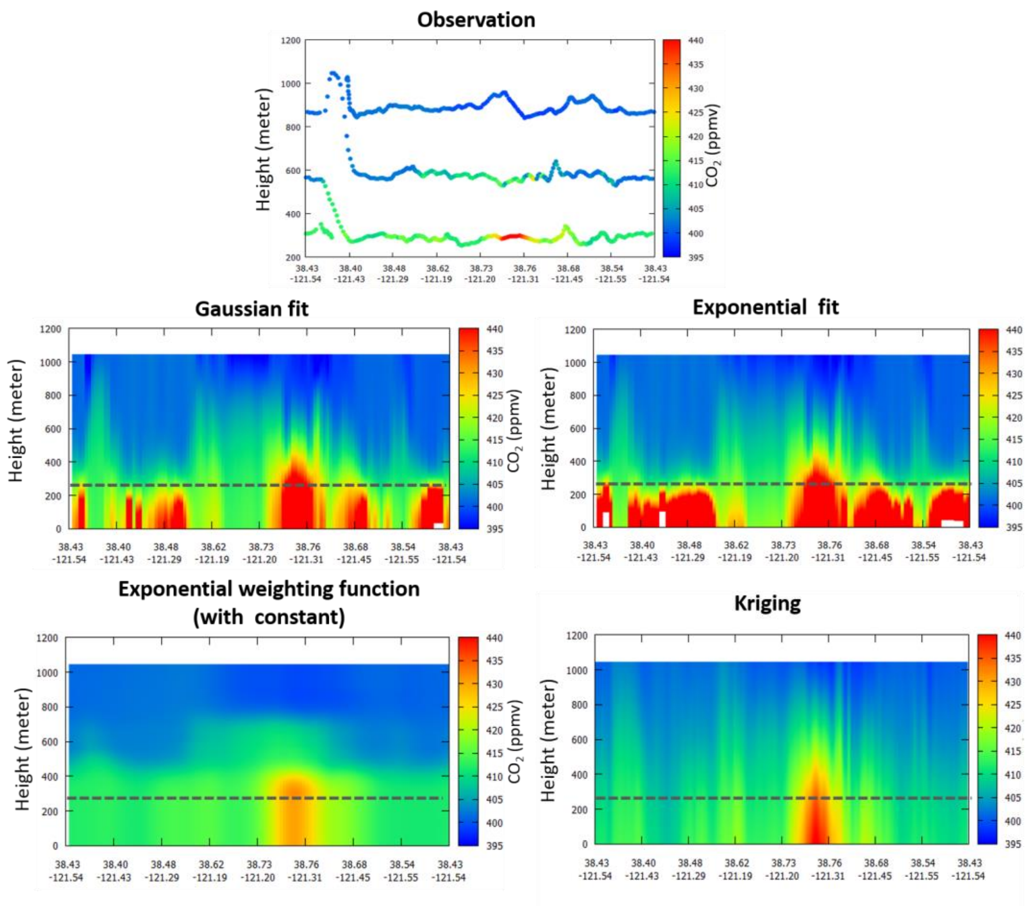
l. 30: Who uses continental scale wind for flux estimates?
We will remove the wind rose plots.

Fig. S3: This figure is not mentioned in the main manuscript. Please add flight dates to the left of the plots.
We mentioned this figure in Line 144 in the original manuscript, but we did not fully explain these plots in the main manuscript. Yes, we will add flight dates to them in the revised version.

l. 36: Consider "falling". For methane the dashed line is blue. Remove "observation".
We will work on them in the revised manuscript.

Fig. S4: Why are there "boxes" or vertical cuts visible in (d)? Does this have to do with gridding? What is the grid size? Could you plot the measurements on top of the interpolated fields? This way it is easier to assess your statement "kriging reflects the individual plume characteristics better". Could you show the extrapolated fields to the ground as well? Which step is performed first: interpolation or extrapolation? Is this described in the text?
Also: Use the same color bar range for all plots.

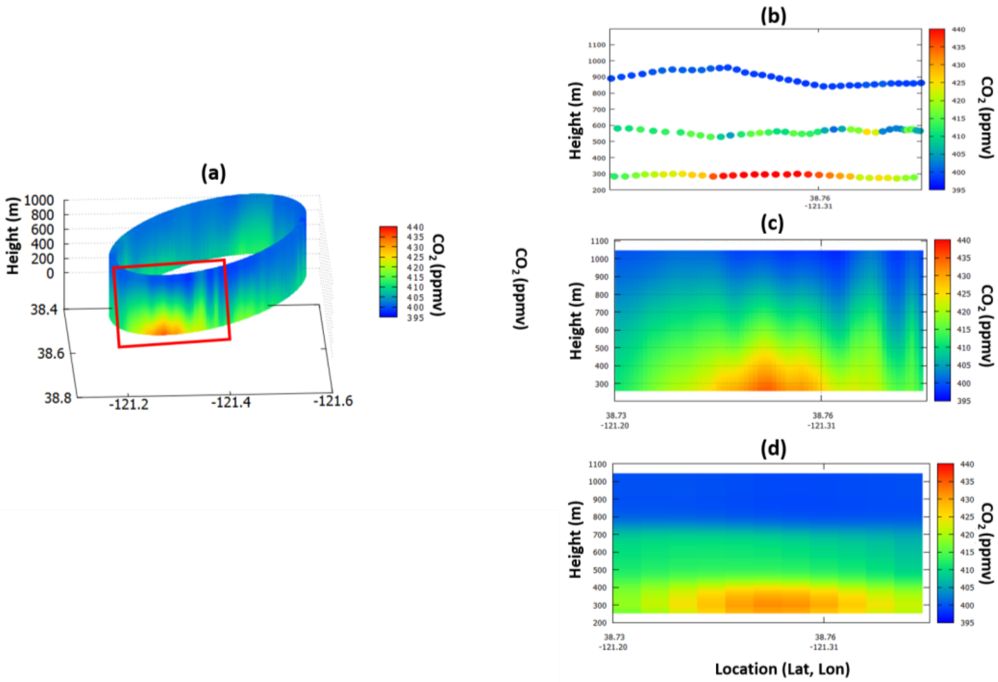
Yes, the boxes (vertical columns) are related to the bin size for the interpolation and fit. Above the lowest flight level, we can use interpolation, but below the flight level, we need to do extrapolation. The white boxes represents no result due to the lack of the number of data used.



The figures above show the CO_2 field extrapolated to the ground. We do both interpolation and extrapolation in one process. We applied a formula for Gaussian Fit and exponential fit (Gordon et al., 2015) based on the lowest flight level data). For the interpolation, we used the exponential weighting function for the data above lowest flight level (for the interpolation), and then a constant value for the locations below the lowest flight level (for the extrapolation). Yes, we will use the same color bar range for all panels. The black dashed line represents an approximate lowest flight line.

l. 57: Don't (b) and (d) also show only the subset of the ellipse? Could you change the direction of these plots? Then this arrow would not be necessary.

The horizontal range is much larger than the vertical range. So, it is very hard to see the actual difference if you try to compare the whole ellipse. That is why we just try to show only a subset of the ellipse for comparison. But as you see in the plots above, there is still a noticeable difference between kriging and interpolation with an exponential weighting function. Yes, I did change the direction of this plot. The modified plot is below.



l. 82: Remove the sentence: “The CH_4 enhancement was localized near the landfill.” This is obvious.

We will remove it in the revised manuscript.

l. 83: Also remove “…, and we … case.” This is also obvious.

We will remove it in the revised manuscript.

Fig. S7: This figure is not mentioned in the main manuscript.

We will discuss it in the main manuscript in the revised version.

Reference

Conley, S., I. Faloona, S. Mehrotra, M. Suard, D. H. Lenschow, C. Sweeney, S. Herndon, S. Schwietzke, G Petron, J. Pifer, E. A. Kort, and R. Schnell: Application of Gauss’s theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, Atmos. Meas. Tech., 10, 3345-3358, https://doi.org/10.5194/amt-10-3345-2017, 2017.

Gordon, M., Li, S.-M., Staebler, R., Darlington, A., Hayden, K., O’Brian, J., and Wolde, M.: Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations, Atmos. Meas. Tech., 8, 3745-3765, doi:10.5194/amt-8-3745-2015, 2015.

Jeong, S., et al.: Estimating methane emissions in California’s urban and rural regions using multi-tower observations, J. Geophys. Res. Atmos., 121, 13,031–13,049, doi:10.1002/2016JD025404, 2016.

Karion, A. et al.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophys. Res. Lett., Vol. 40, 1-5, doi:10.1002/grl.50811, 2013.

Karion, A. et al.: Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region, Environ. Sci. Technol. 2015, 49, 8124-8131, DOI: 10.1021/acs.est.5b00217, 2015.

Kountouris, P., Gerbig, C., Rodenbeck, C., Karstens, U., Koch, T. F., and Heimann, M., 2018: Atmospheric CO_2 inversions on the mesoscale using data-driven prior uncertainties: quantification of the European terrestrial CO_2 fluxes.

Nathan, B., Golston, L. M., O'Brien, A. S., Ross, K., Harrison W. A., Tao, L., Lary, D. J., Johnson, D. R. Covington, A. N., Clark, N. N., and Zondlo, M. A.: Near-Field Characterization of Methane Emission Variability from a Compressor Station Using a Model Aircraft, Environ. Sci. Technol, 2015, 7896-7903, DOI: 10.1021/acs.est.5b00705, 2015.

Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., and AEROCARB contributors, 2005; Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys., 5,2173-3186, 2005.

Turnbull, J. C., Karion, A., Fischer, M. L., Faloona, I., Guilderson, T., Lehman, S. J., Miller, B.R., Miller, J. B., Montzka, S., Sherwood, T., Saripalli, S., Sweeney, C., and Tan, P.P.:: Assessment of fossil fuel carbon dioxide and other anthropogenic trace gas emissions from airborne measurements over Sacramento, California in spring 2009. Atmos. Chem. Phys., 11, 705–721, 2011, doi:10.5194/acp-11-705-2011, 2011.